ICASE

DISCRETE FOURIER TRANSFORM OF VECTOR SETS

AND PERTURBATIONS OF CIRCULANT MATRICES

Igor Najfeld

Report No. 81-19

July 17, 1981

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

DISCRETE FOURIER TRANSFORM OF VECTOR SETS AND PERTURBATIONS OF CIRCULANT MATRICES

Igor Najfeld

Institute for Computer Applications in Science and Engineering

ABSTRACT

The discrete Fourier transform can be adapted to handle vector sets (vectors with set valued components); the result is a vector set. The basic operation is the Minkowski addition. In case of vector balls and vector polygons explicit computation can be carried out. We make use of these results to:

1. calculate exact eigenvalue bounds for a perturbed circulant matrix and
2. derive convergence criterion for an iterative process governed by uncertain circulants.

This research was supported by NASA Contract No. NAS1-15810 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

1. Introduction

The linear transformation of a vector $z \in \mathbb{C}^n$,

$$\hat{z} = Fz \ , \tag{1}$$

where $F \in \mathbb{C}^{n\times n}$ is the Fourier matrix of order n ,

$$F^* = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ \vdots & & & & \vdots \\ 1 & \omega^{n-1} & \omega^{n-2} & \cdots & \omega \end{bmatrix} , \quad \omega = e^{i\frac{2\pi}{n}} , \tag{2}$$

is called discrete Fourier transform (DFT).

Suppose now that z is a set valued vector, that is, each component of z belongs to a bounded set in the complex plane

$$z_k \in Z_k \subset \mathbb{C} \quad , \quad k = 1,\ldots,n \ . \tag{3}$$

Denote such vector set by $\mathbb{Z}$. What is the DFT of $\mathbb{Z}$? An answer to this question depends on the choice of arithmetic operations that deal with sets as operands. Here, we choose the Minkowski addition of sets and the usual scalar multiplication defined as follows: for any two sets A and $B \subset \mathbb{C}^d$ and a scalar $\alpha \in \mathbb{C}$

$$\begin{aligned} A \oplus B &= \{c : c = a + b \, , \, a \in A \, , \, b \in B\} \\ \alpha A &= \{c : c = \alpha a \, , \, a \in A\} \ . \end{aligned} \tag{4}$$

A geometric interpretation of the Minkowski sum is that it represents the set of all possible ordinary sums of a and b when they move independently over set A and set B respectively.

An arbitrary linear combination of sets follows by induction from definition (4) as

$$\begin{aligned} \bigoplus_{k=1}^{n} \alpha_k A_k &= \alpha_1 A_1 \oplus \alpha_2 A_2 \oplus \ldots \oplus \alpha_n A_n \\ &= \{c : c = \sum_{k=1}^{n} \alpha_k a_k \ , \ a_k \in A_k\} \end{aligned} \tag{5}$$

$$A_1,\ldots,A_n \subset \mathbb{C}^d \quad , \quad \alpha_1,\ldots,\alpha_n \in \mathbb{C} \ .$$

Having a definition for linear combinations, it is possible to set up a linear transformation of a vector set. Let $T \in \mathbb{C}^{m\times n}$ be an arbitrary linear transformation and $\mathbb{V}$ a vector set of dimension n,

$$\mathbb{V} = \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix} \quad , \quad V_k \subset \mathbb{C}^d \quad , \quad k = 1,\ldots,n \quad . \tag{6}$$

Then $T \circ \mathbb{V}$ is defined as follows:

$$T \circ \mathbb{V} = \begin{bmatrix} t_{11} V_1 \oplus \ldots \oplus t_{1n} V_n \\ \vdots \\ t_{m1} V_1 \oplus \ldots \oplus t_{mn} V_n \end{bmatrix} = \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix} = \mathbb{U} \quad . \tag{7}$$

Thus, a linear transformation of a vector set, denoted by $\circ$, is the usual linear transformation where the Minkowski addition replaces the ordinary addition in order to fit set valued operands. For more information about $\oplus$ and $\circ$ see for example [2].

Definition: The discrete Fourier transform of a vector set $\mathbb{Z}$ is

$$\hat{\mathbb{Z}} = F \circ \mathbb{Z} \quad . \tag{8}$$

The plausibility of this definition stems from a geometric interpretation; vector set $\hat{\mathbb{Z}}$, which is well defined, represents the set of all possible ordinary DFT's when z varies over $\mathbb{Z}$, that is,

$$\hat{\mathbb{Z}} = \{\hat{z} : \hat{z} = Fz \ , \ z \in \mathbb{Z}\} \quad . \tag{9}$$

From a computational point of view (9) is a poor algorithm even if applied statistically. There is, however, an effective procedure that relies heavily on the geometric structure of sets. The simplest case is that of balls. Let $\mathbb{B}(c,r)$ be a vector ball, that is, each component is a ball in $\mathbb{C}^d$,

$$\mathbb{B}(c,r) = \begin{bmatrix} B(c_1,r_1) \\ B(c_2,r_2) \\ \vdots \\ B(c_n,r_n) \end{bmatrix} \quad , \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad , \quad r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \tag{10}$$

$$c_k \in \mathbb{C}^d \quad , \quad k = 1,\ldots,n \quad .$$

The set of vector balls is closed under the operation $\circ$, and explicit formulae exists for the calculation of centers and radii. Under the linear transformation T, vector ball $\mathbb{B}(c,r)$ is mapped into another vector ball, say $\mathbb{D}$, as follows:

$$\mathbb{D}(\gamma,\rho) = T \circ \mathbb{B}(c,r) = \mathbb{B}(Tc, |T|r) \quad . \tag{11}$$

Here, $|T|$ denotes the matrix whose elements are $|t_{ij}|$, and by the product $\gamma = Tc$ we mean

$$\gamma_k = \sum_{j=1}^{n} t_{kj} c_j \in \mathbb{C}^d \quad , \quad k = 1,\ldots,m \quad . \tag{12}$$

Returning to (8), suppose now that $\mathbb{Z}(c,r)$ is a vector ball of data, that is, each component Z_k is a ball (with center c_k, radius r_k) containing, for example, uncertainty of measurements.

Theorem: The DFT of $\mathbb{Z}(c,r)$ is a vector ball, $\mathbb{Z}(\hat{c},\hat{r})$ with centers

$$\hat{c} = F c \tag{13}$$

and radii

$$\hat{r} = |F|r = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \frac{1}{\sqrt{n}} J r \quad . \tag{14}$$

Proof: Use of definition (8), the closure property (11) and the fact that $|\omega^{j\ell}| = 1$ are the steps in the proof.

An algorithm for fast Fourier transforms (FFT) must be slightly modified in order to deal with (13). Since each $c_k \in \mathbb{C}^d$, the computational complexity of a modified FFT turns out to be $O(nd\,\ell n_2\, n)$

because there are $O(nd)$ essential multiplications and additions per step. On the whole, computation of ball valued Fourier coefficients has the same complexity as (13), since (14) requires n operations.

It is interesting to note that the radii of the ball coefficients are the same, and equal to a scaled arithmetic average of the radii of the ball data. This uniform spread of uncertainty is not accidental if one recalls that the coefficients in the Fourier expansion of a point valued function are equal to certain "rotational" averages $(\int e^{ikx} f(x)dx)$ and that the balls are invariant under unitary transformations ("rotations"). Moreover, since

$$|F| = |F^*| = \frac{1}{\sqrt{n}} J \ , \tag{15}$$

then, squaring both sides of (15) results in

$$|F|^2 = |F^*|^2 = J = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \ ,$$

which can be written as

$$|F| = " \sqrt{J} " \ , \tag{16}$$

thus symbolically exhibiting the least square property of an averaging operation by DFT.

It should be noted that it is impossible to recover the original data set since in general one has

$$\mathbb{B} \subset T^{-1} \circ (T \circ \mathbb{B})$$

and

$$\mathbb{B} \subset T \circ (T^{-1} \circ \mathbb{B})$$

for any nonsingular T and any vector set $\mathbb{B}$ (for proof see [2]).

Suppose now that each Z_k belongs to a plane convex polygon P_k with ν_k vertices; frequently P_k is a real interval (with 2 vertices) or a complex interval (with 4 vertices, see Fig. 1).

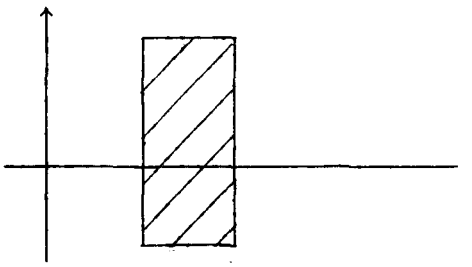

Figure 1

The vector set $\mathbb{P} = (P_1,\ldots,P_n)^T$ is called (plane convex) vector polygon.

Theorem: The class of vector polygons is closed under $\circ$. If $\hat{\mathbb{P}} = F \circ \mathbb{P}$ then the number of vertices for each component of $\hat{\mathbb{P}}$ satisfies

$$\max(\nu_1,\ldots,\nu_n) \leq \hat{\nu}_k \leq \sum_{j=1}^{n} \nu_j \qquad (17)$$

$$k = 1,\ldots,n$$

Proof: The Minkowski sum of two plane convex polygons, Q_1 and Q_2, with μ_1 and μ_2 vertices respectively is a plane convex polygon Q with μ vertices satisfying

$$\max(\mu_1,\mu_2) \leq \mu \leq \mu_1+\mu_2 \qquad (18)$$

For a constructive proof of this statement and an algorithm see [3]. From this and the representation of $\hat{P}_k$,

$$\hat{P}_k = \bigoplus_{j=1}^{u} \omega^{(k-1)(j-1)} P_k \quad , \qquad (19)$$

as a linear combination of n convex polygons, closure, as well as the bounds

in (18), follow by finite induction.

Corollary: If $\mathbb{I}$ is a vector interval then $\mathbb{P} = F \circ \mathbb{I}$ such that each component of $\mathbb{P}$ is a centrally symmetric convex polygon with at most $n(2n)$ vertices if n is even (odd).

Proof: Each interval is a centrally symmetric convex polygon with 2 vertices (by a central symmetry we mean that there exists translation τ of a convex body Γ such that if $\gamma \in \Gamma \oplus \tau$ then $-\gamma \in \Gamma \oplus \tau$, $\forall\ \gamma \in \Gamma$). The linear combination (5) of centrally symmetric convex convex bodies is again a centrally symmetric convex body. Combining these observations and the right hand side of (17) it follows that $2n$ is the maximum attainable number of vertices for each component of $\mathbb{P}$. If n is even, further simplification occurs due to central symmetry of the set of vectors $1, \omega, \ldots, \omega^{n-1}$, i.e., $\omega^k = -\omega^{n/2+k}$ so that only $n/2$ different directions exist. Hence, the maximum attainable number of vertices is $2 \cdot (n/2) = n$.

The next corollary (the proof of which is almost identical to the one above) is motivated by the use of inverse DFT, namely when the transform is a vector complex interval $\mathbb{I}_c$ (see Fig. 1).

Corollary: If $\mathbb{I}_c$ is a vector complex interval then $\mathbb{P} = F^* \circ \mathbb{I}_c$ such that each component of $\mathbb{P}$ is a centrally symmetric convex polygon with at most $2n(4n)$ vertices if n is even (odd).

2. Perturbations of Circulants

The close relationship between DFT and circulant matrices suggests an approach to the perturbation of eigenvalues of circulants. Let

$$A = \text{circ}\ (a_1, a_2, \ldots, a_n) \tag{20}$$

and suppose that each $a_k \in A_k \subset \mathbb{C}$. It is well known that the eigenvalues of circulants are given by

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \sqrt{n}\, F \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad . \tag{21}$$

If a_k varies independently over A_k then, by (8), $\lambda_k \in \Lambda_k \subset \mathbb{C}$ and the explicit, computable, bounds on the eigenvalues of A are

$$\mathbb{\Lambda} = \begin{bmatrix} \Lambda_1 \\ \vdots \\ \Lambda_n \end{bmatrix} = \sqrt{n}\, F \circ \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} = \sqrt{n}\, F \circ \mathbb{A} \quad . \tag{22}$$

3. Iteration with Uncertain Circulants

Consider an iteration process

$$Y_{m+1} = A(m)\, Y_m \quad , \qquad Y_0 \quad \text{given} \tag{23}$$

where $A(m)$ is a fixed matrix picked at the m^{th} stage from

$$A(m) \in \text{Circ}\ (A_1, A_2, \ldots, A_n) \tag{24}$$

where $A_k \subset \mathbb{C}$ are given sets representing uncertainty. This problem was suggested by P. J. Davis who considered a similar problem in [1].

Clearly, convergence (boundedness) of (23) requires that the product

$$P = \prod_{m=1}^{\infty} A(m) \tag{25}$$

converges (be bounded). This question is simplified by the fact that all circulants are diagonalizible by F. Hence,

$$P = F^* \left\{ \prod_{m=1}^{\infty} \text{diag}\,(\lambda_1(m),\ldots,\lambda_n(m)) \right\} F \quad . \tag{26}$$

By the result (22), of the previous section, $\lambda_k \in \Lambda_k$, where

$$\begin{bmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_n \end{bmatrix} = \sqrt{n}\, F \circ \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} \tag{27}$$

so that

$$\Lambda_k = \bigoplus_{j=1}^{n} A_j\, \omega^{(k-1)(j-1)} \quad . \tag{28}$$

It follows that the necessary and sufficient condition for boundedness of (23) is given by

$$\Lambda_k \subset B(0,1) \qquad \text{(unit ball)} \qquad k = 1,\ldots,n \quad . \tag{29}$$

This statement should be interpreted as follows: no matter how one picks sequence $A(m)$ from (24), and/or Y_o, boundedness will result provided condition (29) holds.

Specializing, assume $A_j = B(c_j, r_j)$ is a ball. Then Λ_k is a ball too given by

$$\Lambda(\gamma_k, \rho_k) = \bigoplus_{j=1}^{n} A_j\, \omega^{(k-1)(j-1)} = B\Big(\sum_{j=1}^{n} c_j\, \omega^{(k-1)(j-1)}\,,\ \sum_{j=1}^{n} r_j\Big)$$

that is,

$$\gamma_k = \sum_{j=1}^{n} c_j \, \omega^{(k-1)(j-1)} \tag{30}$$

$$\rho_k = \sum_{j=1}^{n} r_j = \rho \quad . \tag{31}$$

The condition (29) is equivalent to

$$\max_{k = 1,\ldots,n} |\gamma_k| + \rho \leq 1 \quad . \tag{32}$$

Note that $\rho = \sum_{j=1}^{n} r_j \leqslant 1$ is a necessary condition,

while $\sum_{j=1}^{n} |c_j| + r_j \leqslant 1$ is a sufficient condition for boundedness.

References

[1] Philip J. Davis, Circulant Matrices, pp. 99-107, John Wiley & Sons, 1979.

[2] Igor Najfeld, Linear Transformations of Vector Sets and Their Applications, in Interval Mathematics, ed. Karl Nickel, pp. 417-432, Academic Press, 1980.

[3] Igor Najfeld, Analytic Design of Compensators and Computational Geometry, Ph.D. thesis, Brown University, 1978, pp. 54-66.